

学校编码: 10384

分类号_____密级_____

学号: X2013231968

UDC _____

厦门大学

工程硕士学位论文

基于数据仓库的房地产行业 CRM 系统设计 与实现

Design and Implementation of CRM System for Real Estate
Industry Based on Data Warehouse

李建祥

指导教师姓名: 赖永炫副教授

指导教师: 赖永炫副教授

专业名称: 软件工程

论文提交日期: 2016 年 9 月

论文答辩日期: 2016 年 11 月

学位授予日期: 2016 年 12 月

指导教师: _____

答辩委员会主席: _____

2016 年 9 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（ ） 1. 经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

（ ☒ ） 2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

摘 要

随着银行信贷阀门收紧、拿地门槛太高，资金、土地等关系着地产企业发展命脉的核心因素左右着房地产企业的生死沉浮。随着行业的不断变化，许多房地产企业逐渐开始重视企业自身的信息化建设与应用，因为信息化的建设能够帮助企业更好地度过难关。房地产企业的核心内容在于房产的销售，因此许多房地产企业将重心移到了销售系统的信息化管理上，希望通过信息化的手段来维护客户的资源，聚焦客户的价值。因此当前形成了利用 CRM 和数据仓库以及数据挖掘技术来对房地产的客户进行管理,不仅能够从接触客户开始到客户最终购房交付对客户进行管理，并且可以在信息化的管理过程中将客户的信息进行整合，实现客户价值的最大化。

本文在经过对房地产客户的需求以及房地产企业的相关业务信息进行分析 and 了解后，提出了客户信息化管理的模式，并且包含了不同的管理模块，分别为市场研究、会户分析、售楼管理、会员管理以及租赁管理。基于该业务模块，本文设计了结合数据仓库的 CRM 系统平台，该平台由数据仓库采集客户互动中的 Web、客户留言等信息，CRM 系统也相应地包含了以上的各个子系统。在数据方面，本文将数据划分成了前端的数据展示部分、数据存储于管理、数据的分析以及数据的服务提供四个方面，并在数据仓库体系上按照房地产实际业务情况及系统需求，根据以数据驱动的设计原则，确定多个主题。最后基于用例细分，形成了关键数据库 ER 图和数据库表。在数据仓库实现中，重点实现了 ETL 的不同层次的数据加工逻辑，并且针对数据提供方式也制定了应用规则。而在客户分析上，主要通过数据挖掘等一系列算法，对数据仓库中已经汇集的数据信息进行数据挖掘，本文以采用基于 ID3 决策树的决策分析算法，介绍重点客户分析的过程。本系统的实现，完成了房地产企业以客户为中心的 CRM 体系的建立，确保了房地产的营销的智能化及有效性，实现房地产企业的利润最大化和客户服务体验最优化。

关键词：房地产行业；CRM；数据仓库

Abstract

With the tightening of bank credit valve, with threshold is too high, such as capital, land in relation to the real estate enterprise development, the lifeblood of the core factors about the ups and downs of life and death of the real estate business. In the face of drastic changes in the industry, information construction and application are more and more real estate enterprises are looking forward to, and become the core of one of the enterprises through the difficulties. Real estate companies concerned about the focus has become sales management information, continuous operation and maintenance customers and focus on customer value. The formation of the CRM and data warehouse and data mining technology in management of real estate customer philosophy, has formed the management platform based on information system, the effective management of the customer life cycle, customer value maximization.

From the business of real estate customer information management and user requirements, in-depth analysis of real estate customer information management, key business module division to sell building management, lease management, market research, membership management, customer analysis in the paper. Based on the service module is designed in this paper. Combined with the data warehouse of CRM system, the platform by collecting customer interaction of data warehouse in the web, customer information such as mail, and CRM system also divided sale floor management, lease management, membership management, market research management and membership analysis subsystem subsystem. In the design of the data warehouse system, the design of the front end of the display part, a data analysis layer, data storage and management layer and data source provides four layers, and in the data warehouse system according to the real estate business and system requirements, according to the principle of data driven design determined multiple themes; finally based on dividing the case, the formation of the key database ER diagram and database table. In the realization of data warehouse. The focus of the ETL of different levels of data processing logic, and for data provide a way to also formulate rules to

apply; and in the customer analysis, mainly through the data mining algorithms, the data warehouse has been pooled data mining. In this paper, the decision based on the ID3 decision tree analysis algorithm, introduces the process analysis of key customers.

The realization of the system, the completion of the real estate business to customer for CRM system established, to ensure that the intelligence and the effectiveness of real estate marketing, real estate enterprises to maximize profits and customer service experience optimization.

Key Words: Real Estate Industry; CRM; Data Warehouse

目 录

第一章 绪论	1
1.1 研究背景与意义	1
1.2 研究现状	2
1.3 本系统的特点	3
1.4 论文章节安排	3
第二章 相关技术综述	5
2.1 数据仓库	5
2.1.1 概念	5
2.1.2 ETL 技术	6
2.1.3 ETL 流程设计原则	7
2.1.4 常用的数据仓库软件	8
2.2 CDC 技术	8
2.3 数据采集方式	9
2.3.1 源系统接入方式	9
2.3.2 数据采集方式	9
2.3.3 增量采集方法	9
2.3.4 数据采集的基本原则	10
2.4 本章小结	11
第三章 系统需求分析	12
3.1 业务和用户需求分析	12
3.2 功能性需求	12
3.2.1 售楼管理子系统	15
3.2.2 租赁管理子系统	17
3.2.3 会员管理子系统	19
3.2.4 市场研究管理子系统	20
3.2.5 会员分析子系统	21

3.3 非功能性需求	22
3.3.1 系统性能需求.....	22
3.3.2 安全性需求.....	23
3.4 本章小结	23
第四章 系统设计	24
4.1 系统整体设计	24
4.2 功能设计	25
4.2.1 售楼管理子系统.....	26
4.2.2 租赁管理子系统.....	31
4.2.3 会员管理子系统.....	34
4.2.4 市场研究管理子系统.....	35
4.2.5 会员分析子系统.....	37
4.3 数据仓库设计	37
4.3.1 总体架构.....	37
4.3.2 主题设计.....	37
4.3.3 ETL 设计	38
4.3.4 数据采集设计.....	40
4.4 数据库设计	43
4.5 本章小结	51
第五章 系统实现	52
5.1 系统开发环境	52
5.2 数据仓库实现	52
5.2.1 ETL 关键功能实现	52
5.2.2 ETL 作业实现.....	58
5.3 业务系统实现	59
5.4 本章小结	64
第六章 系统测试	65

6.1 系统测试环境	65
6.2 测试用例设计和结果	65
6.2.1 性能测试.....	65
6.2.2 功能测试.....	66
6.2.3 安全性测试.....	71
6.3 本章小结	72
第七章 总结与展望	73
7.1 总结	73
7.2 展望	73
参考文献.....	75
致 谢.....	77

CONTENTS

Chapter 1 Introduction.....	1
1.1 Background and Significance	1
1.2 Study Situation	2
1.3 System Features	3
1.4 Thesis Chapter Arrangement	3
Chapter 2 Summary of Related Technology	5
2.1 Data Warehouse	5
2.1.1 Concepts	5
2.1.2 ETL Technology	6
2.1.3 Common Data Warehouse Software	7
2.2 CDC Technology	7
2.3 Data Collection Mode	8
2.3.1 Source System Access	8
2.3.2 Data Collection Mode	8
2.3.3 Incremental Acquisition Method	8
2.4 Summary	9
Chapter 3 System Requirements Analysis	10
3.1 Business and User Needs Analysis	10
3.2 Functional Requirements	10
3.2.1 Sales Management Subsystem	13
3.2.2 Lease Management Subsystem	14
3.2.3 Membership Management Subsystem	15
3.2.4 Market Research Management Subsystem	15
3.2.5 Member analysis subsystem	16
3.3 Non-Functional Requirements	16
3.3.1 System Performance Requirements	16

3.3.2 Security Requirements	17
3.4 Summary	17
Chapter 4 System Design	18
4.1 Overall SystemDesign	18
4.2 Functional Design	19
4.2.1 Sales Management Subsystem	20
4.2.2 Lease Management Subsystem	27
4.2.3 Membership Management Subsystem	31
4.2.4 Market Research Management Subsystem	34
4.2.5 Member Analysis Subsystem	36
4.3 Data Warehouse Design	36
4.3.1 Overall Architecture	36
4.3.2 Theme Design	37
4.3.3 ETL Design	38
4.3.4 Data Acquisition	41
4.4 Database Design	44
4.5 Summary	52
Chapter 5 System Implementation and Shows	53
5.1 System Development Environment	53
5.2 Data Warehouse Implementation	53
5.2.1 ETL Key Functions to Achieve	53
5.2.2 ETL Job to Achieve	59
5.2.3 Customer Analysis to Achieve	60
5.4 Information System Implementation	61
5.4 Summary	64
Chapter 6 System Test	65
6.1 System Test Environment	65

6.2 Test Case Design and Results	65
6.2.1 Performance Test	65
6.2.2 Function Test	66
6.2.3 Security Test	71
6.3 Summary.....	72
Chapter 7 Summary and Outlook	73
7.1 Summary	73
7.2 Outlook	73
References	75
Acknowledgements	77

第一章 绪论

本章主要讨论房地差 CRM 系统的研究背景、研究意义，在分析研究基础上提出了对于这个系统的研究目标与内容，最后介绍本文的主要组织结构。

1.1 研究背景与意义

在当今复杂的商业环境中，企业想获得竞争优势比以往任何时候都难，甚至于保持优势也变得很难。在快速的全球通信和信息无所不在的时代，具有对客户和市场的深入洞察能力是至关重要的，也是企业获胜的基于。面对客户不断变化的需求和竞争对手不断调整的策略，成功取决于企业是否能够快速地、有准备地以及充满自信地做出响应。房地产企业当前逐渐认识到客户资源在房地产营销及发展中的核心作用，建立并巩固良好的客户关系才是最终的制胜之道。

随着银行信贷阀门收紧、拿地门槛太高，资金、土地等关系着地产企业发展命脉的核心因素左右着房地产企业的生死沉浮。随着行业的不断变化，许多房地产企业逐渐开始重视企业自身的信息化建设与应用，因为信息化的建设能够帮助企业更好地度过难关。房地产商品经营在买卖房地产商品之前，必然会对该商品进行大量的商品信息交换。而对于房地产企业而言，在开发过程中需要许多的信息交流。因此，在这个房地产行业，信息是最重要的因素。目前房地产行业的竞争十分接力，经营模式也逐渐从以产品为中心转向以客户为中心，并且激烈的竞争使得一些发展较小、较慢的房地产企业逐渐被淘汰。房地产企业的核心内容在于房产的销售，因此许多房地产企业将重心移到了销售系统的信息化管理上，希望通过信息化的手段来维护客户的资源，聚焦客户的价值。

另外，一般房地产企业对客户信息的掌握并没有那么全面及时，而是大部分都由企业内的销售人员所掌握。这种发展模式带来的弊端就是，如果该企业的销售人员离职，那么该企业便同时失去了销售人员手上所掌握的客户资料。对于代理模式而言，客户信息则掌握在代理商手上，大部分的销售活动都受到代理商的限制，导致房地产企业不能及时掌握客户资料和动态。例如，在营销方面，经常出现企业的营销活动不能真正满足客户需求的情况，从而给企业带来了很大的损失。因此，房地产企业真正需要的是一套客户关系管理系统，通过网络资源来建立起客户的关系群，并且通过集中式的管理，从而实现企业内

部的不同部门可以通过分享客户资料的方式提高各个部门的工作效率，从而减少因为沟通而带来的费用。另外，企业也能够对客户资料进行大规模的分析和挖掘，从而设计出适合不同客户群体，满足客户需求的营销模式，增加企业的销售收入，提高利润^[1-3]。

对于目前许多的房地产企业而言，他们虽然掌握了较多的客户资源，但是因为客户资源在不同的时间和地点收集，因此大多都被存放在不同的部门和不同的数据库中，这些数据会因为业务系统的不同而存在一些矛盾之处。如果直接将这些数据进行复制来使用，没有通过数据的清洗和集成，这些数据基本上都是无效的数据，因为它们缺乏数据之间的联系性和数据内容的丰富性。对于现有的企业数据库系统，它们的应用环境不适合于对客户信息进行联机分析，也不适用于企业的决策支持。因此，目前形成了通过 CRM、数据挖掘以及数据仓库技术来对房地产的客户信息进行管理的方式^[4-7]，不仅能够从接触客户开始到客户最终购房交付对客户进行管理，而且能够在管理过程中最大化地实现客户的价值。因此，建立一套基于 CRM 的可以用于对房地产客户信息进行全局管理的数据库是非常必要的。

1.2 研究现状

目前国内外已经有许多公司有进行房地产客户关系管理的研究与开发，主要包括 web 技术、.net 技术、J2EE 技术以及基于 B/S 的离线智能 CRM。在数据仓库方面，数据仓库的应用也随着计算机应用的发展而不断衍生，涉及到各行各业的应用，例如互联网金融、保险行业以及零售行业等。数据仓库的应用，使得许多行业在生产销售方面都实现了利益的扩张，从而带来了更多企业对数据仓库的应用和探索。在数据仓库的应用方面，它主要通过与数据挖掘技术来结合，提高数据的并行性和可扩展性，从而与互联网进行集成。另一方面，数据挖掘技术工具的引入还可以将数据仓库用于一系列的打包应用中。在房地产行业，数据仓库的应用面主要包括信息化建设缺乏整体规划，对房地产应用中数据仓库的初始定位不准确。一方面觉得只要在企业中建立了数据仓库，就能够解决企业在客户管理方面的问题。而因为国外的情况与国内不同，许多厂商直接将国外的技术搬到国内来，不仅忽视了数据的分析和使用，而且缺少了系统的集成化，不够成熟，不能够将其直接进行应用，缺乏可移植性。

1.3 本系统的特点

本文经过汲取国内外系统建设经验，分析和设计了基于数据仓库的 CRM 管理理念与信息技术的应用系统。该系统包含以下的子系统，分别为售楼管理、市场研究、会户分析、租赁管理以及会员管理子系统。子系统主要功能有：

（1）售楼管理子系统：报表管理、交易管理、市场营销、客户管理、基础设置、财务管理、售后管理、价格管理；

（2）租赁管理子系统：报表管理、财务管理、市场营销、基础设置、招商管理、客户服务、交易管理；

（3）会员管理子系统：积分管理、事件管理、基础资料、活动管理、报表管理、会员管理；

（4）市场研究管理子系统：地产资讯、项目调查、宏观经济、字典维护、市场报告；

（5）会员分析子系统主要由通过电话、电邮、传真、Web 等数据获取客户信息和产品信息并保存至数据仓库中，并且生成对客户信息的分析内容，包括分析客户的基本信息、未来发展情况、客户忠诚度以及客户实现利润方面的信息，也可以通过数据仓库发现重点的客户资源，包括维持现有的重点客户、发掘潜在客户、对客户进行交叉销售等。基于如上设计，本系统具备的特点有：

（1）鲜明的地产行业特色。设计的系统完全基于房地产行业客户特征、产品特征、销售流程、服务模式、营销手段等，避免了通用的 CRM 的问题。

（2）将先进的房地产营销运营模式进行融合，形成优秀的客户管理模式，包括对客户进行会员管理、提高客户的体验服务等。

（3）先进的系统平台。采用开源的系统框架进行开发 B/S 系统，拥有统一的业务基础平台，支持业务参数、查询、报表、功能等功能处理。

（4）搭建的个性化数据仓库平台，实现多种抽数方案，实现客户的挖掘与分析，强化了系统的深层次营销。

1.4 论文章节安排

本章主要研究内容如下：

1、分析和研究数据仓库的数据组织结构和 ETL (Extract-Transform-Load, 抽取-转换-加载) 的应用机制，结合房地产 CRM 的实际业务需求，设计与实现

数据仓库模型，并做好相关 ETL 策略研究。

2、通过对房地长行业 CRM 系统的总体业务需求分析，提出系统的总体设计思路，按照数据仓库理论对 CRM 系统的框架、功能模块进行设计。

本研究的主要内容包括以下 6 章：

第一章绪论，开展整个论文研究课题内容的相关信息阐述，包括背景，意义，研究现状，技术路线，以及研究内容等等。

第二章关键技术，阐述了数据仓库的一些相关技术，例如 CDC 产品、ETL 技术等。

第三章基于数据仓库的房地产行业 CRM 系统的需求分析，针对与房地产 CRM 的业务需求内容，分析与确定各个功能节点需求内容，以及非功能性需求内容和业务流程内容。

第四章基于数据仓库的房地产行业 CRM 系统系统设计，对目前需求设计的原则性内容进行了分析，并且针对本文系统的整体架构来将系统进行模块划分，从而实现了各个子模块的详细设计；并且详细设计了 ETL 的过程；通过用例图生成数据实体，进而设计数据 ER 图，最终设计出关系数据库表。

第五章基于数据仓库的房地产行业 CRM 系统实现，针对于系统的实现进行展示，配以多种图示；并就数据仓库的 ETL 实现过程进行了阐述，并对客户分析模块以客户重要性分析为例进行介绍。

第六章是总结及展望，一方面将本文所作的工作内容进行了总结，另一方面也提出了工作中存在的一些不足之处，并针对不足之处提出了进一步的展望。

Degree papers are in the “[Xiamen University Electronic Theses and Dissertations Database](#)”.

Fulltexts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.